Gelman and Rubin Diagnostics

Gelman and Rubin diagnostics (Gelman and Rubin; 1992; Brooks and Gelman; 1997) are based on analyzing multiple simulated MCMC chains by comparing the variances within each chain and the variance between chains. Large deviation between these two variances indicates nonconvergence.

Define $\{\theta^t\}$, where $t = 1, \ldots, n$, to be the collection of a single Markov chain output. The parameter θ^t is the tth sample of the Markov chain. For notational simplicity, θ is assumed to be single dimensional in this section.

Suppose you have M parallel MCMC chains that were initialized from various parts of the target distribution. Each chain is of length n (after discarding the burn-in). For each θ^t, the simulations are labeled as θ^t_m, where $t = 1, \ldots, n$ and $m = 1, \ldots, M$. The between-chain variance B and the within-chain variance W are calculated as

$$B = \frac{n}{M-1} \sum_{m=1}^{M} (\bar{\theta}_m - \bar{\theta}_{\cdot\cdot})^2, \quad \text{where } \bar{\theta}_m = \frac{1}{n} \sum_{t=1}^{n} \theta^t_m, \quad \bar{\theta}_{\cdot\cdot} = \frac{1}{M} \sum_{m=1}^{M} \bar{\theta}_m$$

$$W = \frac{1}{M} \sum_{m=1}^{M} s_m^2, \quad \text{where } s_m^2 = \frac{1}{n-1} \sum_{t=1}^{n} (\theta^t_m - \bar{\theta}_m)^2$$

The posterior marginal variance, $\text{var}(\theta \mid y)$, is a weighted average of W and B. The estimate of the variance is

$$\widehat{V} = \frac{n-1}{n} W + \frac{M+1}{nM} B$$

If all M chains have reached the target distribution, this posterior variance estimate should be very close to the within-chain variance W. Therefore, you would expect to see the ratio $\widehat{V}/W$ be close to 1. The square root of this ratio is referred to as the *potential scale reduction factor* (PSRF). A large PSRF indicates that the between-chain variance is substantially greater than the within-chain variance, so that longer simulation is needed. If the PSRF is close to 1, you can conclude that each of the M chains has stabilized, and they are likely to have reached the target distribution.